The Dissertation Committee for Yasuaki Sakamoto
certifies that this is the approved version of the following dissertation:

The Roles of Isolation and Differentiation in Enhanced Oddball Memory

Committee:

Bradley C. Love, Supervisor

W. Todd Maddox

Arthur B. Markman

Risto Miikkulainen

Jane M. Richards

The Roles of Isolation and Differentiation in Enhanced Oddball Memory

by

Yasuaki Sakamoto, B.A.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2005

To my family

Acknowledgments

I wish to express my gratitude to my supervisor, Professor Brad Love, for his continuous support on this project and my committee members, Professors Todd Maddox, Art Markman, Risto Miikkulainen, and Jane Richards for their valuable comments. I am very grateful for the love and support of my wife Momoko, my parents Shinichi and Yoko, my sisters Kumi and Keiko, my parents-in-law Kenichi and Akemi, my sister-in-law Naomi, and my brother-in-law Kenta. This work was financially supported by AFOSR grant FA9550-04-1-0226 and NSF CAREER grant 0349101 to my supervisor.

Yasuaki Sakamoto

The University of Texas at Austin
May 2005

The Roles of Isolation and Differentiation in Enhanced Oddball Memory

Publication No. ____________________

Yasuaki Sakamoto, Ph.D.
The University of Texas at Austin, 2005

Supervisor: Bradley C. Love

What makes a person, event, or object memorable? Enhanced memory for oddball items has been long established, but the basis for these effects is not well understood. This dissertation offers a novel way to think about novelty that clarifies the roles of isolation and differentiation in establishing new memories. According to the isolation account, items that are highly dissimilar to other items are better remembered. In contrast, recent category learning studies suggest that oddball items are better remembered because they must be differentiated from other similar items. The present work pits the differentiation and isolation accounts against each other. The results suggest that differentiation, not isolation, leads to more accurate memory for deviant items. In contrast, gains for isolated items are attributable to reduced confusion with other items, as opposed to preferential storage.

Contents

Chapter 1

1.1 Introduction

Vancouver, Toronto, Montréal, Austin. Given a list of items to remember, people show a memory advantage for an item that differs from the others in some way, such as an American city (Austin) in a list of Canadian cities (Vancouver, Toronto, Montréal). This robust memory phenomenon has been long established and is known as the von Restorff (1933) effect.

The von Restorff effect can be regarded as a specific case of novelty effects. Whether information is deviant depends on how humans structure their environment. In the above example, the structure is the tendency of items in the list to be Canadian cities. Once a structure is discovered, items that are unexpected become novel in the context of the structure. Novelty effects have been established in various forms. For example, deviant faces (Valentine, 1991), behaviors (Hastie & Kumar, 1979), and category members (Palmeri & Nosofsky, 1995) are remembered better than typical items.

Novelty detection is the flip side of stimulus generalization and thus likely plays a central role in our mental development. Indeed, infants tend to show preference for a novel stimulus once they habituate to a familiar one (Fantz, 1964) and this ability to respond to novelty is predictive of later intelligence (McCall & Carriger, 1993). Items judged as novel tend to be processed more fully and deeply than other items and may in fact be processed differently (Friedman, 1979). For instance, people fixate more often and for a longer durations on novel items (Loftus & Mackworth, 1978) and notice changes in the novel items more accurately (Goodman, 1980).

Novelty also plays a major role in social cognition and judgments. For instance, a distinctive member of a group is judged as more influential and more behaviors of the distinctive member are remembered (Taylor, Fiske, Etcoff, & Ruderman, 1978). Moreover, both the positive and negative characteristics of distinctive individuals are perceived as more extreme (McArthur, 1981). Given the importance of novelty in our mental activities, recent work in cognitive neuroscience has focused on identifying the neural circuits underlying novelty processing (e.g., Fabiani & Donchin, 1995; Kishiyama, Yonelinas, & Lazzara, 2004; Ranganath & Rainer, 2003).

Despite the widespread interest in novelty effects, the basis for these effects is not sufficiently well understood. In this dissertation, I pit two accounts of novelty effects against each other. According to the isolation account, items that are highly dissimilar to other items are better remembered. In contrast, the differentiation account holds that items that are highly similar to other items yet differ on some critical properties are better remembered. Previous research has not carefully examined these two accounts.

In the remainder of the paper, I first review relevant work on novelty effects and describe the isolation and differentiation accounts. I then present the results from three experiments that tease apart the contributions of isolation and differentiation to enhanced oddball memory. Finally, I discuss methodological and theoretical implications of the present work to research concerning memory and mental representations.

1.2 Previous Work on Novelty Effects

The current work is guided by the assumption that existing knowledge structures play an important role in determining which items are best remembered. One such candidate structure is a schema. A schema is a general knowledge structure that provides a set of expectations based on prior experience (Brewer & Treyens, 1981; Graesser & Nakamura, 1982; Hastie, 1981; Taylor & Crocker, 1981). For example, a person may have a schema for birds that when activated makes properties like flying and laying eggs available. Schemas can guide the encoding and retrieval of information (Alba & Hasher, 1983; Brewer & Nakamura, 1984; Loftus & Mackworth, 1978; Pichert & Anderson, 1977; Srull, 1981). Category information learned from examples can serve similar functions (Goldstone, 1994; Schyns & Murphy, 1994; Wisniewski & Medin, 1994).

Work in the schema and categorization literatures addresses a related set of issues. I draw parallels between these two literatures and argue that findings from both literatures suggest that items tend to be better remembered to the extent that they conflict with an established knowledge structure (i.e., novelty effects). It might seem odd that such parallels are not already firmly established. One explanation for the disconnect is the varying methodologies and priorities of the two fields.

Work in schemas and stereotypes tends to utilize concepts that are already meaningful to subjects. In contrast, the majority of work in category learning tends to employ artificial categories composed of geometric stimuli that have no meaning outside of the experimental context. In a typical category learning experiment, subjects learn to assign geometric stimuli to one of two mutually exclusive categories (e.g., categories A and B) through trial by trial classification learning with corrective feedback (e.g., Estes, 1994; Maddox & Ashby, 1993; Medin & Schaffer, 1978; Nosofsky, 1988). The category learning work that does involve meaningful prior knowledge tends to focus on how such knowledge can facilitate the acquisition of novel categories (Murphy & Allopenna, 1994; Pazzani, 1991; Wattenmaker, Dewey, Murphy, & Medin, 1986). Accordingly, error rate is the primary dependent measure for the majority of work in category learning, whereas measures of recognition and recall figure more prominently in the schema literature. Nevertheless, work from both areas bears on the research questions considered here.

1.2.1 Relevant and Irrelevant Information

Schema and stereotype research demonstrates that items are encoded to the extent that they are relevant to the activated schema (e.g., Anderson & Pichert, 1978). In Pichert and Anderson's (1977) study, for example, subjects read a story about a house from either a burglar's or a home-buyer's perspective (i.e., schema). The story contained pieces of information that are relevant to one perspective but irrelevant to others. For example, information about a color television set was relevant to the burglar's perspective but irrelevant to the home-buyer's perspective. Alternatively, information about leaking roof was relevant to the home-buyer's perspective but irrelevant to the burglar's perspective. The main finding was that after reading the story subjects better recalled information that was relevant to the perspective they took than irrelevant information.

The ability to selectively encode relevant information is critical as humans are confronted with more information than they can process. Irrelevant information has no obvious connections to the schema and thus is ignored or filtered out. In the present work, novel items are deviant items that are inconsistent with the schema. Inconsistent information is relevant to the schema because it violates the expectations provided by the schema. Thus, relevant information can be further divided into consistent and inconsistent information. For example, encountering a book in a library would be schema-consistent (i.e., in accord with expectations), whereas encountering a concert stage would be schema-inconsistent. Whether schema-consistent or schema-inconsistent information is better remembered is a central issue in schema research.

1.2.2 Consistent and Inconsistent Information

Work in social beliefs and stereotypes has found a memory advantage for schema-consistent information relative to schema-inconsistent information (Rothbart, Evans, & Fulero, 1979; Snyder & Uranowitz, 1978). Social schemas are proposed to function as filtering devices for inconsistent information that lead to inconsistent information being discounted during the encoding process (Taylor & Crocker, 1981). For example, an accountant's rowdy behavior at a party can simply be explained away by inferring the accountant was drunk.

The schema-consistent memory advantage has been challenged by other studies that demonstrate that schema-inconsistent information is remembered better than schema-consistent information (Bower, Black, & Turner, 1979; Goodman, 1980; Hastie & Kumar, 1979; Pezdek, Whetstone, Reynolds, Askari, & Dougherty, 1989). For example, Hastie and Kumar (1979) presented subjects with a list of synonymous adjectives that created a coherent impression of a character. After acquiring this "person schema," behaviors that were inconsistent with this schema were better remembered than those that were consistent.

Rojahn and Pettigrew (1992) conducted a meta-analysis for memory for schema-consistent and schema-inconsistent information and resolved the apparent contradictions across studies. When measures of recognition are corrected for false alarm rate, schema-inconsistent information is remembered better than schema-consistent information. For a library schema, a common false alarm might be reporting to have seen a book when in fact a book did not appear in any studied scene. Stangor and McMillan (1992) conducted a similar meta-analysis in stereotype research and reached the same conclusion as Rojahn and Pettigrew. This tendency to false alarm to consistent information can also be seen in the Deese-Roediger-McDermott false memory paradigm (Deese, 1959; Roediger & McDermott, 1995).

The schema-inconsistent memory advantage is analogous to the memory advantage for novel items in the von Restorff effect (Koffka, 1935; von Restorff, 1933; Wallace, 1965). Unlike typical work in schemas, but like typical work in category learning, subjects gain an appreciation for the structure of the study items during these studies. Once subjects acquire an expectation for the items, the deviant item is analogous to schema-inconsistent information. The von Restorff effect can be seen as a bridge between work in schema research that relies on pre-existing knowledge structures and work in category learning in which expectations are only developed after a number of learning trials.

1.2.3 Separate Organization of Novel Items

Experiments in the tradition of von Restorff and schema research indicate a memory advantage for items that conflicts with an established structure or expectation over items that conform to the structure. The deviant items are novel in the context of the other items once the structure or expectation is discovered. What makes novel items more memorable? Earlier accounts focused on differential attention allocated to novel items at the time of encoding (Jenkins & Postman, 1948; Green, 1956). For instance, Green proposed that the change from preceding items leads to increase in attention allocated to novel items. According to this account, novel items result in enhanced memory because increased attention to the novel items induces additional processing of those items.

However, the differential attention account is at odds with more recent work that demonstrates that deviant items are remembered better even when presented at the beginning of a study list (Dunlosky, Hunt, & Clark, 2000; Hunt & Lamb, 2001). Because no structure has been acquired at the time the novel items are presented, there will be no additional attention and processing for the novel items at the time of encoding. Thus, more recent work focuses on the separate organization of novel items (Bruce & Gaines, 1976; Fabiani & Donchin, 1995; Hunt & Lamb, 2001). According to this idea, a deviant item can become distinctive by grouping of other items subsequent to the presentation of the deviant item. The separate storage of novel items result in enhanced memory for those items.

1.2.4 Isolation vs. Differentiation

Most explanations along the line of separate organization focus on the advantage conferred to items that are relatively isolated from other stimulus items (e.g., Busey & Tunnicliff, 1999; Hunt & Lamb, 2001; Schmidt, 1991). According to the isolation account, novel items are better remembered when they are more separated (i.e., isolated) from the other items (e.g., Busey & Tunnicliff, 1999; Schmidt, 1991; Valentine, 1991). The basic idea is that the most dissimilar item occupies an isolated region of a similarity space, is the least confusable, and results in best memory.

As I review in the following section, however, recent category learning research that affords greater experimental control has brought the isolation account into question and has instead suggested that the basis for the memory advantage of novel items is contrasting with highly similar items that establish a context or backdrop (Sakamoto & Love, 2004). The differentiation account focuses on how specific or highly tuned memories for novel items are. Items that are contrasted with highly similar items have more opportunities for confusion with other items, which should lead to more differentiation and finer-grained memory representations.

These two accounts have not been distinguished in previous research. One reason is that whether isolation or differentiation account is operable depends on how the relationships between the novel and other items are construed. In most situations, oddball items are not only isolated but also differentiated.

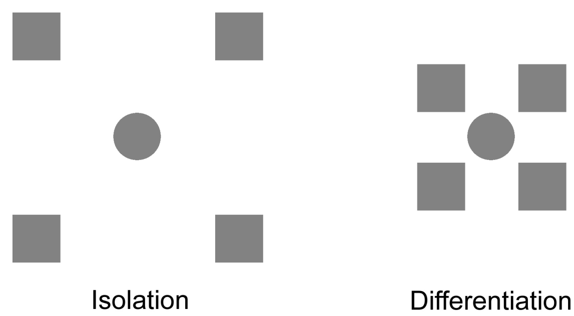

Figure 1.1: Examples of isolation and differentiation are shown. The left circle is more isolated, whereas the right circle is more differentiated.

In Figure 1.1, for example, both circles are isolated as they have a unique property (i.e., shape) that the other items do not have. Both circles are also differentiated as they share properties with the other items (e.g., color, size) but deviate on the shape. However, the two circles differ in their degrees of isolation and differentiation. The left circle is more isolated because it is farther away from the surrounding squares than the right circle. In contrast, the right circle is more differentiated as it is closer to the surrounding squares than the left circle.

Inter-item similarity relations play opposing roles in the isolation and differentiation accounts. In the isolation account, items that are highly dissimilar and atypical are best remembered. In contrast, in the differentiation account, items that are highly similar to other items, yet deviate on a critical property (such as category membership or shape in the case of Figure 1.1), are best remembered.

1.2.5 Related Work in Category Learning

Whereas the isolation account is the main explanation of novelty effects in the memory literature (Busey & Tunnicliff, 1999; Valentine, 1991; Schmidt, 1991), findings from category learning studies suggest that deviant items may be better remembered because they must be differentiated from other similar items. In Palmeri and Nosofsky's (1995) Experiment 3, for example, subjects learned to classify geometric stimuli into two contrasting categories. An imperfect regularity successfully classified the majority of study items (e.g., most small items are in Category A, whereas most large items are in Category B), but two exceptions violated the regularity (e.g., a large item that is a member of Category A). Once subjects learn the category structure, the exceptions are analogous to schema-inconsistent items. Following learning, subjects showed a recognition advantage for the exceptions over the rule-following items as in the schema and basic memory research.

Of importance, both the exceptions and the rule-following items in the Palmeri and Nosofsky study shared the same similarity relations with the other items. Thus, the exceptions were not dissimilar to the other items but were better remembered. This suggests that the memory advantage for the exceptions was due to differentiation, not isolation.

Sakamoto and Love's (2004) Experiment 1 relates to pitting isolation and differentiation against each other. Sakamoto and Love modified Palmeri and Nosofsky's design by introducing an asymmetry in the category structures in which one category contained more rule-following items than the contrasting category. In terms of the present work, the differentiated exception that violated the more salient (i.e., more frequent) regularity was highly similar to and had more opportunities for confusion with members of the opposing category, which should lead to a finer-grained memory representation. In contrast, the isolated exception that violated the less frequent regularity was relatively dissimilar to and less confusable with other items from the opposing category. As predicted by the differentiation account, subjects remembered better the differentiated exception than the isolated exception after learning.

In accord with Sakamoto and Love's results, schema and memory research indicate that the strength of the regularity modulates the advantage for deviant items. For instance, Koffka (1935) reported that when there were more anomalous items in a list, the memory advantage for those items was smaller. Similarly, Rojahn and Pettigrew's (1992) meta-analysis suggests that the memory advantage for the schema-inconsistent items becomes weaker as the proportion of the schema-inconsistent items becomes larger, though the effect was not universal. For example, Pezdek, Whetstone, Reynolds, Askari, and Dougherty (1989) found that the proportion of inconsistent items had no effect on memory for inconsistent items. One possible explanation for null effects in schema research is that schemas are well learned prior to the experiment and therefore may not be as sensitive to the frequency manipulations experienced in brief laboratory studies.

1.3 Current Studies

The present work utilizes a category learning procedure to evaluate the relative contributions of isolation and differentiation to enhanced oddball memory. In three experiments, I examine how isolated and differentiated oddball items are represented in memory. Although Sakamoto and Love's results are in favor of the differentiation account, similarity between exceptions and rule-following items was not directly manipulated in their study. The differentiated exception was contrasted not only with highly similar items but also with a larger number of rule-following items. Likewise, schema studies that manipulate the proportion of inconsistent to consistent items do not control inter-item similarity relations. The present work deals with this issue by controlling inter-item similarity relations.

In Experiments 1–3, subjects learned to correctly assign stimuli varying in color (red or green) and length (continuously valued) to one of two contrasting categories through trial by trial classification learning with corrective feedback. As displayed in Figure 2.1, Figure 2.3, and Figure 2.5, the membership of all but two items can be correctly determined by applying an imperfect rule (e.g., red items are in Category A, whereas green items are in Category B). One exception (i.e., oddball) item (e.g., a green line in Category A) in each category violates this regularity.

In Experiments 1 and 2, inter-item similarity relations are manipulated such that the isolated exceptions are dissimilar to other items, whereas the differentiated oddball items are contrasted with highly similar items. To foreshadow the results, qualitatively different memory advantages are attributable to isolation and differentiation manipulations. Consistent with the differentiation account, the differentiated items result in finer-grained item representations than the isolated items. As predicted by the isolation account, the isolated items are easier to identify than the differentiated items.

In Experiment 3, the differentiated and the isolated exception items share the same inter-item similarity relations, but the differentiated item is more frequently contrasted with highly similar items whereas the isolated item is more frequently contrasted with highly dissimilar items. As in Experiments 1 and 2, the differentiated item is more accurately remembered than the isolated exception. Unlike Experiments 1 and 2, however, Experiment 3 demonstrates that the isolated item is no longer easier to identify than the differentiated exception when they are equally separated from the other items.

Chapter 2

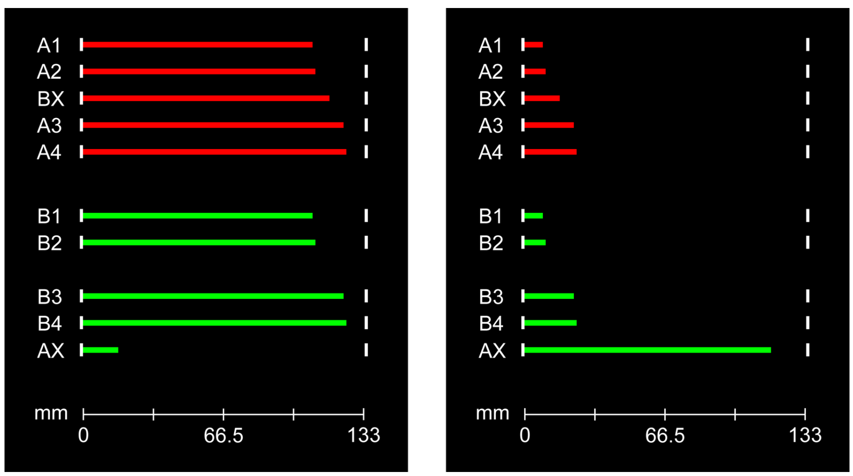

Figure 2.1: The stimuli used in Experiment 1 are shown. Membership in Category A or B can be correctly determined for all but two items by applying an imperfect rule. In this case, red items tend to be in Category A, whereas green items tend to be in Category B. Each category contains an exception. Item BX is more differentiated than item AX, whereas item AX is more isolated than item BX. To eliminate possible influences of absolute line length on performance (Ono, 1967), subjects were randomly assigned to either the left condition in which the differentiated exception was longer than the isolated exception or to the right condition in which the isolated exception was the longer item.

2.1 Experiment 1

In Experiment 1, the differentiated exception (labeled BX in Figure 2.1) was highly similar to items belonging to the contrasting category, whereas the isolated exception (labeled AX in Figure 2.1) was dissimilar to other items. The differentiation account predicts that subjects will develop high-fidelity memory traces for the exception that was more differentiated to reduce confusions with similar items from the opposing category. In contrast, the isolation account predicts that memory should be best for the isolated exception due to its dissimilarity to other items. Memory measures following the learning phase are used to evaluate these accounts of enhanced oddball memory.

2.1.1 Method

Subjects

Seventy-eight University of Texas undergraduates participated for course credit.

Materials, design, and procedure

The experiment was run on Pentium III computers operating in Windows 95. The monitors had 15 inch CRT color displays and a refresh rate of 16.67 ms. The stimuli were simple lines that varied in their color (red or green) and length (continuous) as displayed in Figure 2.1. Instructions were displayed on the monitor at the start of each phase. The background color was black.

Subjects completed a learning phase consisting of classification learning trials of 10 line stimuli. Subjects were provided with the imperfect rule because our main interest was their memory for the exceptions (cf. Palmeri & Nosofsky, 1995; Sakamoto & Love, 2004). The exceptions were manipulated in a within-subjects design such that one exception was highly similar to other items and more differentiated (labeled BX in Figure 2.1), whereas the other exception was highly dissimilar to other items and more isolated (labeled AX in Figure 2.1). To eliminate possible influences of absolute line length on performance (Ono, 1967), subjects were randomly assigned to either the left condition in Figure 2.1 in which the differentiated exception was longer than the isolated exception or to the right condition in Figure 2.1 in which the isolated exception was the longer item. The learning phase ended when subjects completed either 20 blocks of learning trials or two consecutive error free blocks, whichever occurred first. A block is the presentation of each learning item in a random order.

On each trial in the learning phase, one stimulus was randomly positioned around the center of the monitor. The text "Category A or B?" and an imperfect rule was presented above the stimulus. For example, "If the line is red, then Category A. If the line is green, then Category B" appeared below the "Category A or B?" text. The instruction stated that this strategy may not work all the time and that there are two exceptions. Subjects indicated their category membership judgment by pressing the A or B key. After responding, the text and the rule above the stimulus were replaced with visual (e.g., "Right! The correct answer is A.", "Wrong! The correct answer is B.") and auditory corrective feedback (i.e., a low-pitch tone for errors and a high-pitch tone for correct responses). The stimulus and the visual feedback were displayed for 2000 ms after responding. The feedback for the exceptions were twice as long as than that for the rule-following items and included the text "This is one of the exceptions." to facilitate learning. Then, a blank screen was displayed for 2000 ms and the next trial began.

After completing the learning phase, subjects completed a filler phase consisting of three arithmetic problems to prevent rehearsal of information from the learning phase. Each problem consisted of two integers (randomly generated between 10 and 49) presented side by side (e.g., 22 + 34 = ?) and the problem remained displayed until the subjects responded. Subjects received both auditory and visual feedback indicating whether they added the numbers correctly.

Then, subjects completed a reconstruction phase, in which they reconstructed the lengths of the differentiated and isolated exceptions from the learning phase. Subjects reconstructed each exception three times, alternating between each on successive trials. The reconstruction phase measures how accurately the exceptions are remembered.

On each trial in the reconstruction phase, an exception was randomly positioned around the center of the monitor on each trial. The exception's initial length was set 66.5 mm, amid the two exceptions' actual lengths. The color of the exception was given and subjects were told that this was the exception. For example, the text "This exception was in Category A." was displayed above the stimuli. Subjects used the less than or the greater than keys to change the line length. Subjects pressed the Z key when the line length was in the desired position. After pressing the Z key, the text "Thank you" was presented beneath the stimulus and a high-pitch tone sounded. The stimulus and the text were displayed for another 2000 ms. A blank screen was displayed for 2000 ms and the next trial began.

After the reconstruction phase, subjects completed another set of filler phase described above. Finally, subjects completed a transfer phase in which they classified the 10 items presented in the learning phase without corrective feedback. Subjects completed 2 transfer blocks. The transfer phase allows for evaluation of subjects' ability to identify or recognize the exceptions through their performance on the two exceptions in the absence of supervised learning. The procedure for the transfer phase was similar to that for the learning phase except that no rule or feedback was provided. After responding A or B, a high-pitch tone sounded and the text "Thank you" was displayed below the stimulus.

2.1.2 Results and Discussion

Table 2.1: Mean accuracies are shown for the learning and transfer phases of Experiment 1. Mean absolute differences (in mm) between subjects' predicted length and the actual length are shown for the reconstruction phase. These absolute differences are used in the t-test analysis. Item types included the isolated exception and the differentiated exception.

Item type	Learning	Reconstruction	Transfer
Isolated	.82	6.1	.90
Differentiated	.43	2.5	.78

Table 2.1 displays subjects' mean performances in the learning, reconstruction, and transfer phases. Four subjects were unable to meet the criterion before completing 20 blocks. As predicted, in the learning phase, subjects classified the isolated exception more accurately (.82 vs. .43) than the differentiated item, $t(77) = 15.04$, $p < .001$. The differentiated exception, which was surrounded by highly similar items, was more difficult to master than the isolated exception.

Reconstruction error was measured as the absolute difference between subjects' predicted length and the actual length ($|\text{Error}| = |\text{predicted length} - \text{actual length}|$). Consistent with the differentiation account, the mean reconstruction error (averaged across three trials) for the differentiated exception was significantly smaller (2.5 mm vs. 6.1 mm) than that for the isolated exception, $t(77) = 6.64$, $p < .001$. Figure 2.2 displays the relative frequency distribution of subjects' reconstruction responses for the differentiated and isolated exceptions. More responses centered around the actual value (i.e., difference of 0) for the differentiated exception. These results suggest that subjects developed more accurate representations for the differentiated exception than the isolated exception.

As predicted by the isolation account, subjects were better able to identify the isolated exception in the transfer classification phase following reconstruction. Although subjects reconstructed the differentiated exception more accurately, their

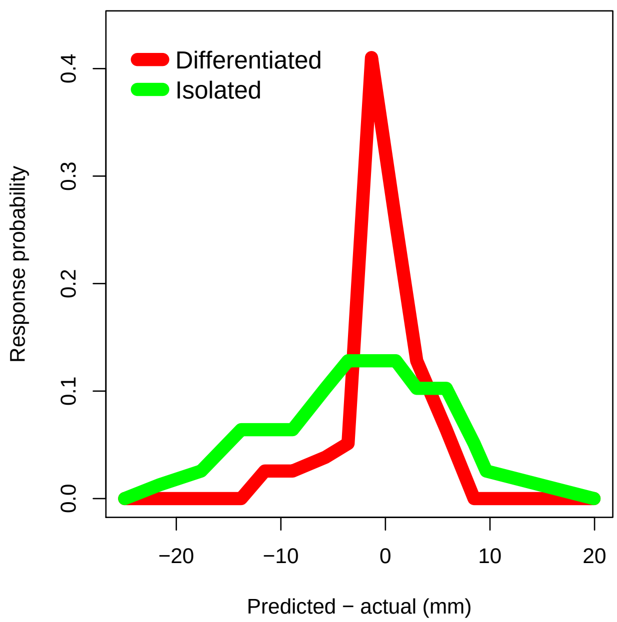

Figure 2.2: Probability distribution of subjects' responses are shown for the differentiated and isolated exceptions in the reconstruction phase of Experiment 1. The x-axis represents the difference in millimeters between subjects' predicted length and the actual length. Positive values indicate overshoot (i.e., predicted length − actual length is greater than zero), whereas negative values indicate undershoot. In the t-test analysis, these difference values are converted into absolute values.

transfer classification performance was significantly better (.90 vs. .78) for the isolated than for the differentiated exception, $t(77) = 3.19$, $p < .01$. The isolated item was easier to identify because it was highly dissimilar to other items and less confusable, as indicated by its relatively high accuracy in the learning phase. Because of less confusability, however, the isolated item results in a lower fidelity memory trace as shown in Figure 2.2. This suggests that the transfer classification advantage for the isolated item will be eliminated if foil items are introduced that are similar to the isolated item (cf. Davidenko & Ramscar, 2004).

Table 2.2: Mean accuracies (and standard errors) are shown for the isolated and differentiated exceptions in the transfer phases of Experiment 1 as a function of reconstruction accuracies. Subjects were sorted from smallest to largest absolute reconstruction errors ($|\text{Error}| = |\text{predicted length} - \text{actual length}|$) and divided into three groups. The high reconstruction accuracy group consisted of the first third, the medium accuracy group consisted of the next third, and the low accuracy group consisted of the last third. For the differentiated exception, a more accurate reconstruction results in a larger space or absolute difference between the reconstructed length and the length of the nearest confusable neighbor form the opposing category ($|\text{predicted length} - \text{nearest confusable neighbor's length}|$). For the isolated exception, a more extreme reconstruction (i.e., the reconstructed length was longer than the actual length when the isolated exception was a long line or shorter than the actual length when it was a short line) leads to more space because the isolated exception was either the shortest or longest line. The isolated exception was highly dissimilar (93.1 mm to the nearest neighbor on one side), resulting in much larger space overall between the reconstructed length and the nearest neighbor's length than the differentiated exception (6.65 mm to the nearest neighbor on each side).

Item type	Reconstruction accuracy (|Error|)	Space	Transfer accuracy
	High (1.92)	91.18	.90 (.05)
Isolated	Medium (4.96)	88.14	.92 (.05)
	Low (11.44)	81.66	.88 (.06)
	High (.84)	5.81	.92 (.04)
Differentiated	Medium (1.80)	4.85	.79 (.06)
	Low (4.99)	1.66	.62 (.08)

Individual analyses

The results from Experiment 1 suggest that more accurate memory does not always result in better identification. I propose that identification is largely determined by how isolated an item is in a representational space rather than the memory accuracy itself. More accurate memory does not necessarily lead to less confusability. An item is less confusable and easier to identify when it is more separated from its nearest confusable item. According to this account, the differentiated exception was not easy to identify because although it was represented more accurately, its representation was still not as separated from the nearest confusable neighbor from the opposing category as the representation of the isolated exception was. In Experiment 1, the isolated exception was highly dissimilar (93.1 mm to the nearest neighbor), resulting in much larger space or absolute difference between the reconstructed length and the length of the nearest confusable neighbor from the opposing category (|predicted length − nearest neighbor's actual length|) than the differentiated exception (6.65 mm to the nearest neighbor on each side).

For the differentiated exception, a more accurate reconstruction results in a larger space between reconstructed length and nearest confusable neighbor's length. If more space leads to better identification, then subjects' reconstruction accuracies (i.e., absolute reconstruction errors) should correlate with their ability to identify the differentiated exception in the transfer classification. As predicted, there was a significant negative correlation ($r = -.26, t(76) = -2.39, p < .05$) between individual subjects' reconstruction errors and transfer classification accuracies. Smaller reconstruction errors on the differentiated exception were correlated with higher classification accuracies of that item in the transfer phase, suggesting that more space between the reconstructed length of the differentiated exception and the actual length of the nearest confusable neighbor is associated with better identification of the differentiated exception.

For the isolated exception, a more extreme reconstruction (i.e., the reconstructed length was longer than the actual length when the isolated exception was a long line or shorter than the actual length when the isolated exception was a short line) leads to more space between reconstructed length and nearest confusable neighbor's length because the isolated exception was either the shortest or longest line. Nevertheless, more accurate reconstructions tended to result in more space because there were few extreme reconstruction responses. There was no significant correlation ($r = -.06, t(76) = -.51, p \approx .61$) between the absolute reconstruction errors and the transfer classification accuracies for the isolated exception. One reason for the null correlation for the isolated exception is that because the isolated exception was highly dissimilar, there was a large space between the predicted length of the isolated exception and the actual length of the nearest confusable item even when the reconstruction was not so accurate. Thus, the isolated exception was easy to identify regardless of the reconstruction errors.

To further test the account that more space between predicted length and the nearest confusable neighbor's length results in easier identification, subjects were divided into three groups according to how accurately they reconstructed the exceptions in the reconstruction phase. Subjects were sorted from smallest to largest absolute reconstruction errors (|predicted length − actual length|) and divided into three groups. The high reconstruction accuracy group consisted of the first third, the medium accuracy group consisted of the next third, and the low accuracy group consisted of the last third.

As displayed in Table 2.2, for the isolated exception, the high, medium, and low reconstruction accuracy groups all had large space between predicted length and the nearest confusable neighbor's length because the isolated exception was extremely dissimilar in Experiment 1. The three groups performed well on the isolated exception in the transfer classification phase, and there were no statistically significant differences among groups. For the differentiated exception, the high reconstruction accuracy group performed better (.92 vs. .62) in the transfer classification of the differentiated exception than the low accuracy group, $t(50) = 3.51$, $p < .01$, with the medium accuracy group in the middle (.79).

The pattern of results generally follow the idea that an item is better identified when its representation is more separated from the nearest confusable neighbor. The isolated exception resulted in large space between predicted length and the nearest confusable neighbor's length and was easy to identify. The differentiated exception was identified more easily when its representation was more separated from the nearest confusable neighbor from the opposing category (see Table 2.2). Further, the same pattern was obtained when subjects were divided into three groups according to the transfer classification accuracy. In general, subjects who were more accurate in the transfer classification showed reconstruction responses that were more separated from the nearest neighbor from the opposing category.

Summary of Experiment 1

Experiment 1 evaluated two accounts of enhanced memory for oddball items by pitting them against each other. According to the differentiation account, items that are highly similar to other items yet differ on some critical properties are better remembered. In contrast, the isolation account holds that items that are highly dissimilar to other items are better remembered. The differentiated exception was more accurately reconstructed as predicted by the differentiation account, whereas the isolated exception was better identified in the transfer classification phase as predicted by the isolation account. The individual analyses suggested that the degree of separation or isolation in a representational space, rather than memory accuracy itself, plays a major role in determining identification performance. How separated an item is from surrounding items in a representational space, inferred from subjects' reconstruction performances, was able to predict their identification performances and vice versa, suggesting that the same representations can account for both more accurate memory for the differentiated exception and easier identification of the isolated exception.

2.2 Experiment 2

In Experiment 1, the isolated item was not only the most dissimilar item but also the most extreme item by virtue of being the shortest or longest line depending on condition. In Experiment 2, both the differentiated and isolated exceptions are

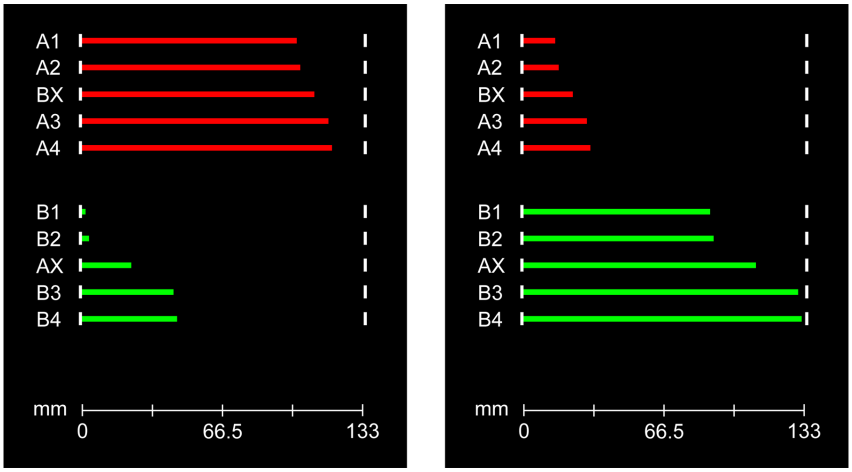

Figure 2.3: The stimuli used in Experiment 2 are shown. As in Experiment 1, two exceptions (one relatively isolated, one relatively differentiated) violated an imperfect rule and subjects were randomly assigned to either the left condition in which the differentiated exception (BX) was longer than the isolated exception (AX) or to the right condition in which the isolated exception was the longer item.

surrounded by other items. As shown in Figure 2.3, the similarity relations between these exceptions and their neighbors are manipulated such that the differentiated exception is more confusable with near members of the contrasting category. Other than this change in category structure, Experiment 2 is identical to Experiment 1.

2.2.1 Method

Subjects

Eighty-two University of Texas undergraduates participated for course credit.

2.2.2 Results and Discussion

Table 2.3 displays subjects' mean performances in the learning, reconstruction, and transfer phases. One subject was unable to meet the criterion of two consecutive error free learning blocks before completing 20 blocks. The main results mirrored

Table 2.3: Mean accuracies are shown for the learning and transfer phases of Experiment 2. Mean absolute differences (in mm) between subjects' predicted length and the actual length are shown for the reconstruction phase. These absolute differences are used in the t-test analysis. Item types included the isolated exception and the differentiated exception.

Item type	Learning	Reconstruction	Transfer
Isolated	.66	2.7	.91
Differentiated	.60	2.1	.77

those of Experiment 1.

Subjects classified the isolated exception more accurately (.66 vs. .60) than the differentiated one in the learning phase, $t(81) = 2.89$, $p < .01$. Following learning, the differentiated exception resulted in smaller absolute reconstruction errors (2.1 mm vs. 2.7 mm) than the isolated one, $t(81) = 3.37$, $p < .01$. More reconstruction responses centered around the actual value for the differentiated exception as displayed in Figure 2.4. Without corrective feedback, the isolated exception, despite being less accurately remembered, was more accurately classified (.91 vs. .77) than the differentiated one, $t(81) = 3.89$, $p < .001$.

Individual analyses in Experiment 2 revealed that, as in Experiment 1, an exception was easier to identify when its reconstructed length was more separated from the actual length of the nearest confusable item from the opposing category. In accord with the idea that more space between the predicted length and the nearest neighbor's actual length leads to better identification, there was a significant negative correlation ($r = -.27, t(80) = -2.55, p < .05$) between individual subjects' reconstruction errors and transfer classification accuracies for the differentiated exception, indicating that smaller reconstruction errors on the differentiated exception (and thus larger space between the predicted length and the actual length of the nearest neighbor's) were associated with better identification of that item. For the isolated exception, there was no correlation ($r = 0$) between the absolute recon-

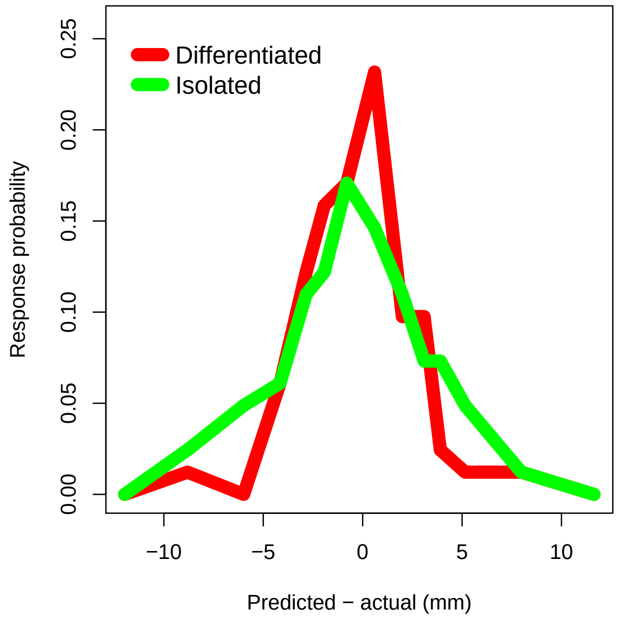

Figure 2.4: Probability distribution of subjects' responses are shown for the differentiated and isolated exceptions in the reconstruction phase of Experiment 2. The x-axis represents the difference in millimeters between subjects' predicted length and the actual length. Positive values indicate overshoot (i.e., predicted length − actual length is greater than zero), whereas negative values indicate undershoot. In the t-test analysis, these difference values are converted into absolute values.

struction errors and the transfer classification accuracies. There was relatively large space between the reconstructed length of the isolated exception and the actual length of the nearest confusable item regardless of the reconstruction errors because the isolated exception was highly dissimilar to the confusable neighbors from the opposing category.

As in Experiment 1, subjects were divided into three groups according to how accurately they reconstructed the exceptions. Table 2.4 displays that, for the isolated exception, the high, medium, and low reconstruction accuracy groups all had relatively large space between the reconstructed length and the nearest neighbor's actual length, and the three groups performed well on the isolated exception in

Table 2.4: Mean accuracies (and standard errors) are shown for the isolated and differentiated exceptions in the transfer phases of Experiment 2 as a function of reconstruction accuracies. Subjects were sorted from smallest to largest absolute reconstruction errors ($|\text{Error}| = |\text{predicted length} - \text{actual length}|$) and divided into three groups. The high accuracy group consisted of the first third, the medium accuracy group consisted of the next third, and the low accuracy group consisted of the last third. For both the isolated and differentiated exceptions, a more accurate reconstruction results in a larger space or difference between the reconstructed length and the length of the nearest confusable neighbor ($|\text{predicted length} - \text{nearest confusable neighbor's length}|$). The isolated exception was highly dissimilar (19.95 mm to the nearest neighbor on each side), resulting in larger space overall between the reconstructed length and the nearest neighbor's actual length than the differentiated exception (6.65 mm to the nearest neighbor on each side).

Item type	Reconstruction accuracy (|Error|)	Space	Transfer accuracy
	High (1.05)	18.90	.94 (.03)
Isolated	Medium (2.35)	17.60	.91 (.04)
	Low (4.64)	15.31	.88 (.05)
	High (.92)	5.73	.85 (.05)
Differentiated	Medium (1.81)	4.84	.80 (.06)
	Low (3.55)	3.10	.66 (.07)

the transfer classification phase with no statistically significant differences among groups. For the differentiated exception, the high reconstruction accuracy group was more accurate (.85 vs. .66) in the transfer classification of the differentiated exception than the low accuracy group, $t(53) = 2.21$, $p < .05$, with the medium accuracy group in the middle (.80). The same pattern was obtained when subjects were divided into three groups according to their transfer classification accuracies. Subjects who were more accurate in the transfer classification tended to have more space between their reconstructions and the nearest confusable neighbors. Thus, more space between the reconstructed length and the nearest neighbor's actual length in general resulted in better identification. Like Experiment 1's results, these results indicate that the same representations underlie memory accuracy and identification ability.

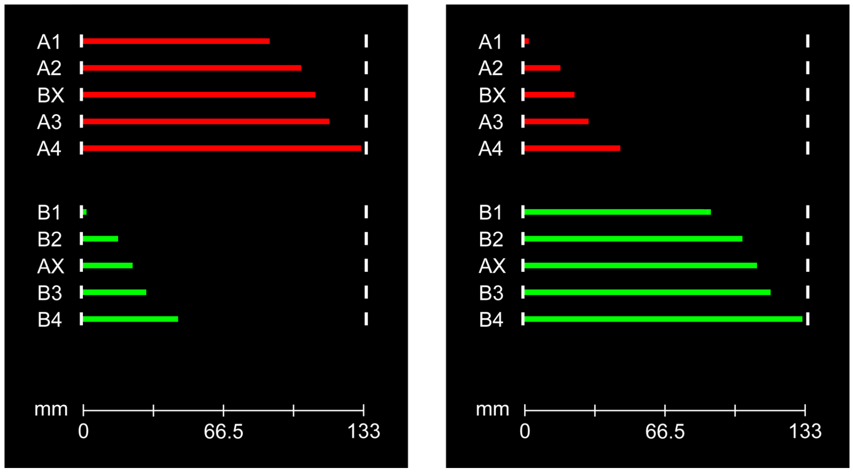

Figure 2.5: The stimuli used in Experiment 3 are shown. Unlike Experiments 1 and 2, the differentiated (BX) and isolated (AX) exceptions shared the same inter-item similarity relations. The differentiated exception was more "differentiated" because its near members of the contrasting category (A2 and A3) were presented more frequently than its distant neighbors (A1 and A4) in the learning phase. The isolated exception was more "isolated" because its distant members of the contrasting category (B1 and B4) were presented more often than its near neighbors (B2 and B3).

2.3 Experiment 3

Experiments 1 and 2 manipulated inter-item similarity relations and demonstrated that items that are highly differentiated from similar items result in more highly-tuned memory representations, whereas items that are highly separated from other items are less confusable and easier to identify. In Experiment 3, similarity relations are equated for the isolated and differentiated exceptions as shown in Figure 2.5. Instead, the differentiated item is more frequently contrasted with highly similar items by presenting its adjacent rule-following items from the opposing category (A2 and A3 in Figure 2.5) more frequently (90% vs. 10%) than its distant rule-following items (A1 and A4) during learning. In contrast, the isolated item is more frequently contrasted with highly dissimilar items by presenting its distant rule-following items from the opposing category (B1 and B4) more frequently (90% vs. 10%) than its adjacent rule-following items (B2 and B3). Other than these changes in similarity relations and presentation frequency of the rule-following items, Experiment 3 is identical to Experiments 1 and 2.

2.3.1 Method

Subjects

Fifty-two University of Texas undergraduates participated for course credit.

2.3.2 Results and Discussion

Table 2.5: Mean accuracies are shown for the learning and transfer phases of Experiment 3. Mean absolute differences (in mm) between subjects' predicted length and the actual length are shown for the reconstruction phase. These absolute differences are used in the t-test analysis. Item types included the isolated exception and the differentiated exception.

Item type	Learning	Reconstruction	Transfer
Isolated	.51	4.2	.81
Differentiated	.44	2.6	.80

Subjects' mean performances in the learning, reconstruction, and transfer phases are displayed in Table 2.5. Eight subjects were unable to meet the criterion of two consecutive error free learning blocks before completing 20 blocks. As in Experiments 1 and 2, the differentiated exception had more opportunities for confusion with members of the opposing category and resulted in a finer-grained memory representation. Subjects classified the isolated exception more accurately (.51 vs. .44) than the differentiated one in the learning phase, $t(52) = 3.32$, $p < .01$. Following learning, the differentiated exception resulted in smaller absolute reconstruction errors (2.6 mm vs. 4.2 mm) than the isolated one, $t(52) = 2.62$, $p < .05$. More reconstruction responses centered around the actual value for the differentiated exception

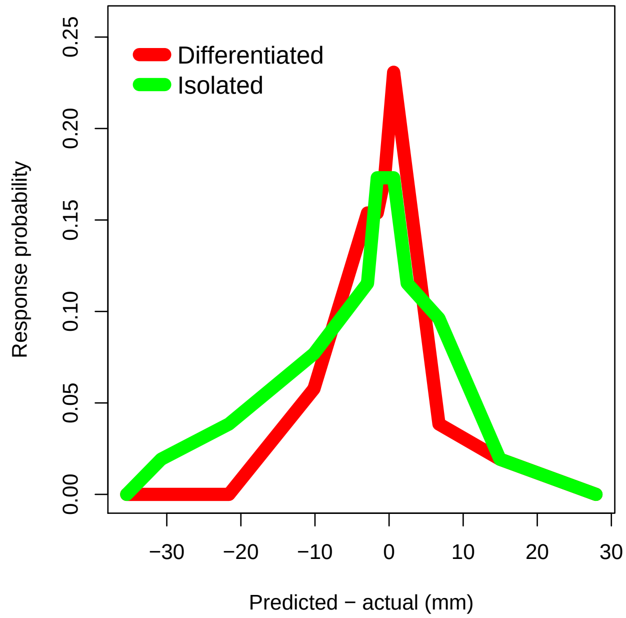

Figure 2.6: Probability distribution of subjects' responses are shown for the differentiated and isolated exceptions in the reconstruction phase of Experiment 3. The x-axis represents the difference in millimeters between subjects' predicted length and the actual length. Positive values indicate overshoot (i.e., predicted length − actual length is greater than zero), whereas negative values indicate undershoot. In the t-test analysis, these difference values are converted into absolute values.

as displayed in Figure 2.6.

Unlike Experiments 1 and 2, the identification advantage for the isolated exception was eliminated when the isolated exception shared the same inter-item similarity relations as the differentiated exception in Experiment 3. Subjects' transfer classification accuracy without corrective feedback for the isolated (.81) and the differentiated (.80) exceptions showed no significant difference, $t < 1$. Although the isolated exception resulted in fewer classification errors than the differentiated exception during learning, they were both highly confusable with other items as indicated by their low accuracies in the learning phase. These results suggest that the frequency manipulation in Experiment 3 plays an important role in memory accuracies but has a minimal impact on identification abilities. The ease of identification is largely determined by the inter-item similarity relations.

Table 2.6: Mean accuracies (and standard errors) are shown for the isolated and differentiated exceptions in the transfer phases of Experiment 3 as a function of reconstruction accuracies. Subjects were sorted from smallest to largest absolute reconstruction errors ($|\text{Error}| = |\text{predicted length} - \text{actual length}|$) and divided into three groups. The high accuracy group consisted of the first third, the medium accuracy group consisted of the next third, and the low accuracy group consisted of the last third. For both the isolated and differentiated exceptions, a more accurate reconstruction results in a larger space or difference between the reconstructed length and the length of the nearest confusable neighbor ($|\text{predicted length} - \text{nearest confusable neighbor's length}|$). The isolated and differentiated exceptions were both 6.65 mm away from the nearest neighbor on each side. A star ($\star$) indicates that the mean absolute error was longer than 6.65 mm, indicating that the predicted length on average was not within the gap between the actual lengths of the exception and the nearest confusable item.

Item type	Reconstruction accuracy (|Error|)	Space	Transfer accuracy
	High (.81)	5.84	.82 (.07)
Isolated	Medium (1.94)	4.71	.97 (.03)
	Low (9.51$\star$)	2.86	.64 (.11)
	High (.61)	6.04	.97 (.03)
Differentiated	Medium (1.81)	4.84	.79 (.09)
	Low (5.31)	1.34	.64 (.10)

Further analyses of individual performance revealed that, as in Experiments 1 and 2, there was a significant negative correlation ($r = -.46, t(50) = -3.66, p < .01$) between individual subjects' reconstruction errors and transfer classification for the differentiated exception. Smaller reconstruction errors on the differentiated exception were associated with better identification of that item, suggesting that larger space between the reconstructed length and the nearest neighbor's actual length leads to better identification. Unlike Experiments 1 and 2, the same pattern was found for the isolated exception ($r = -.63, t(50) = -5.77, p < .01$) in Experiment 3. The isolated and differentiated exceptions shared the same similarity relations with other items. Consequently, both exceptions resulted in similar space between the predicted length and the nearest confusable item's actual length in Experiment 3. This was not the case for the isolated exceptions in Experiments 1 and 2 as those items were more dissimilar to other items than the differentiated exceptions were.

Subjects were divided into three groups according to how accurately they reconstructed the exceptions. Unlike in Experiments 1 and 2, the high, medium, and low reconstruction accuracy groups for the isolated exception had relatively small space between the reconstructed length and the actual length of the nearest neighbor, as displayed in Table 2.6. Although the higher transfer classification accuracy for the medium than high reconstruction accuracy group approached significance ($t(32) = 1.86$, $p \approx .07$), the significant negative correlation for the isolated exception suggests that there was an overall trend that the transfer classification accuracy for the isolated exception was generally better when its reconstruction was more separated from the nearest confusable neighbor. For example, the middle reconstruction accuracy group performed better (.97 vs. .64) in the transfer classification of the isolated exception than the low accuracy group, $t(33) = 2.95$, $p < .01$.

As in Experiments 1 and 2, for the differentiated exception, the high reconstruction accuracy group performed better (.97 vs. .64) in the transfer classification of the differentiated exception than the low accuracy group, $t(33) = 3.18$, $p < .01$, with the medium accuracy group in the middle (.79). Like in Experiments 1 and 2, grouping subjects according to the transfer classification accuracies on the exceptions resulted in the same patter of results for both the differentiated and isolated exceptions. Subjects with more accurate transfer classification tended to have more space between their reconstructions and the nearest confusable neighbors. Taken together, the results from Experiment 3 strongly indicate that the identification ability is largely determined by how separated an item is from other confusable items in a representational space, and that a single representation underlies both memory accuracy and identification.

Chapter 3

3.1 General Discussion

Experiments 1–3 evaluated two explanations for enhanced memory for oddball items. The isolation account holds that deviant items should be remembered best because they are dissimilar to other items. In contrast, the differentiation account holds that items should be remembered best that are similar to other items, yet differ in some critical property (e.g., category membership) that brings the contrast into focus. In everyday life and in most laboratory experiments, both of these explanations are usually operable and their relative influences are indeterminate. For example, in the first example of this paper, Austin is isolated in that it has unique properties that Vancouver, Toronto, Montréal do not possess, but it is also differentiated in that it shares many properties with these other cities, yet varies in the key property of nationhood. Which explanation is operable depends on how one construes the list of city names. To address this conundrum, Experiments 1–3 pitted isolation and differentiation accounts against each other in a well-controlled classification learning task. The studies' designs utilized rule-plus-exception category structures in which one category's exception was relatively isolated and the other's was relatively differentiated.

The results from Experiments 1 and 2 show that isolation and differentiation lead to qualitatively different memory enhancements. As predicted by the differentiation account, the differentiated exception led to a finer-grained item representation than did the isolated exception. As predicted by the isolation account, the isolated exception was easier to identify (i.e., classify as an exception) than the differentiated exception.

Experiment 3's results suggest that the role of frequency in enhanced oddball memory is mediated by similarity. As predicted by the differentiation account, the differentiated exception that was contrasted with similar items more frequently resulted in a finer-grained memory trace than the isolated exception that was contrasted with dissimilar items more often. When the inter-item similarity relations were equated for the isolated and differentiated exceptions, the identification advantage for the isolated exception was eliminated. The results from Experiment 3, coupled with those from Experiments 1 and 2, suggest that whereas the memory specificity is affected by both frequency and similarity information, the identification ability is mostly determined by the inter-item similarity relations.

Indeed, in Experiments 1–3, subjects whose reconstruction of the oddball item was more separated from (i.e., more dissimilar to) the nearest confusable neighbor from the opposing category showed better transfer classification performance on that item (and vice versa). Thus, how isolated an item is in a representational space, as opposed to memory accuracy itself, is a major determinant of whether the item is better identified. Further, the individual analyses demonstrated that subjects' reconstruction responses for the exceptions could predict their identification performance on those items in the subsequent transfer classification phase, indicating that the same representations underlie memory accuracy and identification.

Taken together, Experiments 1–3 demonstrate that novel items that are contrasted with highly similar items that establish a prevailing context result in more accurate memory. Items that are dissimilar to other items lead to easier identification to the extent that those items are highly separated from other confusable items in a representational space. These different types of memory enhancement can be predicted by assuming the same mental representations.

One alternative view is that the number of errors during the learning phase drove these differences as subjects in all experiments made more errors classifying the differentiated exception than the isolated exception during the learning phase. Indeed, similarity and confusability are the catalysts of differentiation and also beget classification errors. However, errors and differentiation are not synonymous. By manipulating the feedback associated with an item, Sakamoto and Love (2004) dissociated violating a regularity and committing an error during learning, and found that enhanced memory is attributable to structure violation and not errors per se.

3.1.1 Methodological Implications

The present results may help resolve the apparent conflict between studies that do and do not find isolation advantages. In Experiments 1 and 2's reconstruction task, the differentiated exception was more accurately remembered, whereas in the transfer phase, the isolated exception was more accurately classified. The isolation advantage was eliminated in Experiment 3 when the isolated and the differentiated exceptions shared the same inter-item similarity relations and were both highly confusable. The isolation advantage in transfer classification of Experiments 1 and 2 can be attributable to the isolated exception being less confusable with members of the opposing category than the differentiated exception is. Thus, the isolation advantage in transfer is likely due to the nature of the other test items, rather than due to a stronger memory trace for the isolated exception.

Analogously, studies that have found an isolation advantage in old/new recognition judgments did not include foils that are similar to isolated items (e.g., Busey & Tunnicliff, 1999; Valentine, 1991). Studies that include foils equally similar to all studied items do not find an isolation advantage (e.g., Davidenko & Ramscar, 2004; Shiffrin, Huber, & Marinelli, 1995; Zaki & Nosofsky, 2001; though see Nosofsky & Zaki, 2003). In contrast, the advantage for the differentiated exception in Experiments 1 and 2's reconstruction task is not attributable to other items included in the test set and instead indicates that subjects developed finer-grained representations for the differentiated exception.

Overall, evidence for the isolation advantage (beyond advantages attributable to reduced confusion with foil items) is limited. In fact, some related findings allude to an isolation disadvantage. For example, humans find average or prototypical items to be most familiar (Posner & Keele, 1968). Further, humans find face images created by averaging numerous images of actual faces to be the most attractive (Langlois, Roggman, & Musselman, 1994). This aesthetic preference for prototypical stimuli extends across a number of domains and has been related to ease of processing or fluency (Rolf, Schwarz, & Winkielman, 2004).

As demonstrated in the present studies, one determinant of whether an isolation advantage or disadvantage is observed is the nature of the task. For instance, item confusability constrains performance for tasks that yield an isolation advantage, whereas confusability is not harmful or even beneficial for tasks not favoring isolation. Future work that employs multiple memory measures and distinguishes between isolation and differentiation will be necessary to fully resolve these issues.

3.1.2 Theoretical Implications

Novelty-gated storage

A number of category learning and memory models utilize novelty detection mechanisms to gate storage (Metcalfe, 1993; Nosofsky et al., 1994; Love, Medin, & Gureckis, 2004). For example, Love et al.'s SUSTAIN clustering model of category learning groups together similar items and forms new clusters in memory in response to surprising events, such as learning that a bat is a mammal and not a bird. This mechanism allows SUSTAIN to correctly predict enhanced recognition memory for stimulus items that violate salient regularities as observed in Palmeri and Nosofsky (1995). SUSTAIN develops rule-following clusters and shifts attention to the rule dimension. When an exception item elicits a prediction error (i.e., surprising event), SUSTAIN recruits an additional cluster to encode the item. While rule-following items tend to cluster with one another, each exception item will be isolated in its own cluster. This differential storage makes exceptions more distinctive in memory.

Similarly, Nosofsky et. al's RULEX hypothesis-testing model correctly predicts enhanced memory for exceptions by explicitly storing items that violate inferred rules. RULEX constructs rules and stores exceptions to the rules. Rule-following items are not individually stored but rather are captured by the rule. Information about exceptions is explicitly stored. The separate storage of exception information allows RULEX to predict the memory advantage of exceptions.

RULEX's and SUSTAIN's treatment of rule-following and rule-violating items parallels findings from the schema research about how consistent and inconsistent information is processed. In particular, RULEX and SUSTAIN are in accord with findings that suggest that people process schema-inconsistent information more deeply and at a greater level of detail. Friedman (1979), for example, demonstrated that missing features, new features, or physical changes in the schema-consistent items were rarely noticed, whereas these changes in the schema-inconsistent items were almost always noticed. Friedman's (1979) proposal that items are only processed to the degree they violate a schema is in accord with SUSTAIN's and RULEX's differential storage of items that violate the regularity. Similarly, Goodman (1980) reported that changes in unexpected items were detected more accurately than changes in expected items. Other proposals unfold along similar lines (Graesser, 1981; Heider, 1946; Schank & Abelson, 1977; Sentis & Burnstein, 1979; Sherman & Frost, 2000).

Clusters vs. rules

To tease apart the predictions of cluster- and rule-based accounts of category representation and to test the differentiation hypothesis, Sakamoto and Love (2004) manipulated the frequency of rule-following items such that one category contained more rule-following items than the contrasting category. As in the present studies, the differentiated exception violating the more salient (i.e., more frequent) regularity had more opportunities for confusion with members of the opposing category than the isolated exception violating the less frequent regularity, and the differentiated exception led to enhanced memory. This finding was predicted by SUSTAIN, but could not be accounted for by RULEX. Rules are insensitive to frequency information and according to RULEX, both the differentiated and isolated exceptions violate the regularities with the same strength.

SUSTAIN predicts enhanced memory for the differentiated exception because SUSTAIN's cluster recruitment is sensitive to frequency of items. A prediction error occurs when SUSTAIN attempts to cluster together highly similar items from competing categories. The exception clusters brought about such errors by attracting rule-following items from the opposing category. There were more opportunities for such errors involving the differentiated exception to occur because the differentiated exception was similar to many rule-following items from the opposing category. Consequently, a greater number of rule-following clusters were recruited that are similar to the differentiated exception, which formed a highly contrastive backdrop for the differentiated exception and enhanced recognition.

However, it is unclear how SUSTAIN's solutions map onto the reconstruction task. Love (2002) presented a clustering model related to the SUSTAIN model that accounted for Palmeri and Nosofsky's (1995) results. Like SUSTAIN, the clustering model stored rule-violating items in their own cluster. Importantly, the model developed sharper tunings (related to memory specificity or distinctiveness) for the cluster encoding the exceptions than the tunings of the cluster encoding the rule-following items (Love, 2002).

The dynamics that drive this outcome are consistent with the explanation of the current results. Each cluster's tuning is adjusted on each learning trial in order to minimize prediction errors. The cluster encoding the differentiated exception tends to be activated by the presentation of rule-following items that match on the rule-relevant dimension (i.e., more confusable items). To avoid these unwanted intrusions, the cluster becomes highly tuned and specific, which minimizes activation by items other than the differentiated exception. The increased distinctiveness of the cluster leads to high fidelity memory trace and enhances its memory accuracy. The same dynamics govern the cluster encoding the isolated exception, but this cluster does not become as distinct as the cluster encoding the differentiated exception because of the similarity/frequency manipulation (i.e., fewer trials in which its tuning is sharpened). Thus, the differentiated exception should result in more accurate reconstruction than the isolated exception. In contrast, the isolated exception should be easier to identify because its cluster is more separated from other clusters encoding the rule-following items from the opposing category.

Similarity-based processing

SUSTAIN (and the clustering model) and RULEX utilize novelty-gated storage and account for enhanced oddball memory because oddball items are stored separately from the other items. However, enhanced memory for the differentiated exception is at odds with rule-based mental representations of regularities on the definition that rules are insensitive to similarity and frequency information (e.g., Pinker, 1991). Of course, additional mechanisms, such as supplementing rules with exemplar storage and weighting the rules by the frequency and/or similarity of the rule-following items, could allow rule-based accounts to be more sophisticated and flexible. When there are no constraints on rules, however, the need for rules is questionable. The current results, coupled with Sakamoto and Love's, favor a clustering account that is more schematic in nature and engages in similarity-based processing. These results suggest that models, as humans, need to be sensitive to both novelty (or structure violation) and similarity information.

In fact, exemplar-based models, which utilize similarity-based processing but lack novelty-gated storage, cannot account for the oddball recognition advantage. Exemplar models store every training instance in memory rather than using novelty gated storage and do not accord special status to oddball items. To determine recognition strength, exemplar models sum the similarity of the probe item to all exemplars stored in memory, which does not predict an advantage for oddball items.

Likewise, exemplar models do not favor items that are odd by virtue of being isolated because the summed similarity recognition calculation favors familiar or typical items. For this reason, when applied to identification tasks, correct identification is modeled as the inverse of summed similarity, thus favoring isolated items (Zaki & Nosofsky, 2001). The idea is that the most dissimilar item is the least confusable and results in the best memory (Busey & Tunnicliff, 1999). Of course, this account is at odds with the current findings suggesting finer-grained representations for differentiated than for isolated exceptions.

An exemplar model with novelty-gated storage

The critical problem with exemplar models is that storage is not dependent on what other items are already stored in memory as in models in which storage is gated by novelty. To support this claim, ALCOVE (Kruschke, 1992), an exemplar model, can account for enhanced oddball memory when novelty-gated storage is incorporated using exemplar-specific attention[1] in which each exemplar has its own attention

[1] see Kruschke, 2001 for a related model with exemplar-specific "specificity"

(i.e., exemplar-specific attention) along a dimension (Sakamoto, Matsuka, & Love, 2004). The idea is that attention is specific to the region along a dimension in the representational space.

Standard exemplar models utilize the same attention at all locations along a dimension (i.e., dimension-wide attention) in the representational space (e.g., Kruschke, 1992; Love et al., 2004; Nosofsky, 1986. The dimension-wide attention is well suited for many artificial category learning studies, in which categories are symmetric and all category members are differentiated by the values on the same dimensions. However, some laboratory work does suggest that humans attend to different dimensions of an item depending on the context the item is in (Aha & Goldstone, 1992; Barsalou & Medin, 1986; Lewandowsky, Kalish, & Ngang, 2002). Such flexible attention is needed when categories contain inconsistent members.

In ALCOVE with exemplar-specific attention (ES-ALCOVE), each exemplar selected which dimensions to attend to. Sakamoto, Matsuka, and Love (2004) demonstrated that in ES-ALCOVE, attention was allocated to the non-rule dimensions of exemplars encoding exceptions but to the rule dimension of exemplars encoding rule-following items. This differential attention made exceptions distinctive in memory.

However, ES-ALCOVE was unable to account for the enhanced memory for the differentiated exception found by Sakamoto and Love (2004). This failure arises because the attention shift in ES-ALCOVE does not distinguish the differentiated and isolated exceptions. The differentiated exception violating the more frequent regularity has more opportunities for confusion with members of the opposing category than the isolated exception violating the less frequent regularity. The differentiated exception results in larger errors or discrepancies between target and predicted output values. ES-ALCOVE treats large and small discrepancies in the same manner and does not distinguish the two exceptions.

Sakamoto, Matsuka, and Love (2004) created another version of ALCOVE called ESSW-ALCOVE for Exemplar-Specific Squeaky Wheel ALCOVE. In addition to the exemplar-specific attention, ESSW-ALCOVE has a mechanism that emphasizes larger errors and minimizes the impact of smaller ones. With accentuated errors[2], ESSW-ALCOVE distinguished important errors (e.g., miss-classification) from trivial ones (e.g., correct classification with 90% confidence level). ESSW-ALCOVE learned attention more rapidly in response to larger errors and minimized the impact of smaller errors. Consequently, ESSW-ALCOVE distributed more attention to the non-rule dimensions of the differentiated exception, which results in larger errors, than to the non-rule dimensions of the isolated exception. Like SUSTAIN and the clustering account, ESSW-ALCOVE better recognized the differentiated item that were more confusable.

ESSW-ALCOVE may be applied to the reconstruction task in the current work. When attention does not sum to a particular number, the differentiated exception should lead to more attention overall in addition to more uniform attention to the non-rule dimensions than the isolated exception because of the larger errors associated with the differentiated exception. More attention for each dimension of the differentiated exception can be interpreted as higher-fidelity memory for that item. Like the clustering model, ESSW-ALCOVE should be able to predict more accurate reconstruction of the differentiated exception than the isolated exception if attention does not sum to a particular number. Using the inverse of summed similarity for identification, the isolated exception should be better identified by ESSW-ALCOVE because it is more dissimilar to other items. It should be noted that ESSW-ALCOVE is analogous to SUSTAIN and the clustering model in that exemplars are grouped together or separated by the attention structures, which can

[2]Errors between target and predicted value was raised to the 10th power instead of the standard sum-squared errors. A similar effect can be obtained by updating the attention weights multiple times on each training trial (e.g., Kruschke, 2001).

be viewed as distributed clusters.

Mental representations

Consideration of the current results, coupled with those from Sakamoto and Love (2004), strongly favor non-rule-based representations of regularities or patterns. Factors such as frequency, expectation violation, and similarity to other items are crucial factors driving performance in these tasks, suggesting that storage is gated by novelty and mental representations are cluster- or schema-like and are engaged through similarity-based processing. Rule-based accounts utilize novelty-gated storage but cannot account for the enhanced memory for the differentiated exception because rules are insensitive to similarity and frequency information. Exemplar-based accounts are sensitive to similarity relations but need novelty detection mechanisms to predict memory advantage for oddball items. Cluster-based representations possess both novelty-gated storage and similarity-based processing, and do a good job accounting for the results described in the current paper.

3.2 Future Work

One concern in the category learning work that examines memory for exceptions is that two components are associated with violating a structure. One component is that the oddball item initially violates the feature expected from the opposing category. That is, the oddball item does not belong to the category it is supposed to (e.g., most small items belong to Category A but the small oddball item belongs to Category B). The other component is that the oddball item violates the regularity of its own category. That is, the oddball item belongs to the wrong category (e.g., most Category B items are large except for the small oddball item). My interpretation focused on initial feature violations rather than later category violations because classification learning procedures encourage people to focus on the diagnostic features (Markman & Ross, 2003; Yamauchi & Markman, 1998) and use rules such as “if small, then Category A.” Thus, the mapping is likely to be from features to categories and the oddball items violate the features.

However, it is unclear whether memory advantage for exceptions is due to initial feature violations or later category violations. It might seem odd that exceptions are better remembered because it violates the feature of the contrasting category and might make more sense if exceptions are distinctive in the context of its own category. Indeed, schema and memory research tends to utilize a single category and deviant items are better remembered because those items are different from other items in the same category.

One way to test whether exceptions are better remembered when they violate the category regularities of their own categories is to use non-diagnostic features for the exceptions. For example, the regularities could be Category A members tend to be small, whereas Category B members tend to be large. The Category A and B exceptions have medium size and do not initially violate the features of the opposing category. These exceptions will be simply different from other items in their own category as in the schema and memory studies.

An interesting question would be if memory advantage for a differentiated exception is observed under this condition. If memory for exceptions in category learning studies is due entirely to initial feature violations, then enhanced memory for the differentiated exception should be eliminated when it only violates the later category regularity. In contrast, schema and memory research suggests that enhanced memory for differentiated exception should be observed even when exceptions do not initially violate the feature regularity of the opposing category.

One possibility is that when a single category is used as in schema and memory work, people attempt to resolve the inconsistency and contrast the deviant item with other members of the category (Srull & Wyer, 1989). The differentiation account predicts that memory for items should be enhanced to the extent that those items are contrasted with highly similar items. If this is the case, then the differentiated exception will be the distinctive item in its own category when exceptions only violate the later category regularities. Future work along this line would advance our understanding of determinants of enhanced oddball memory and increase the connections between schema and category learning research.

3.3 Final Note

Novelty effects have been examined in various domains, including the study of schemas (Rojahn & Pettigrew, 1992), stereotypes (Stangor & McMillan, 1992), basic memory phenomenon (Hunt & Lamb, 2001), face recognition (Valentine, 1991), the neurobiological basis of memory (Kishiyama et al., 2004), and category learning (Sakamoto & Love, 2004). The relative contributions of isolation and differentiation are not well understood in these studies. The present work offers a novel way to think about novelty that clarifies the roles of isolation and differentiation in establishing new memories.

References

Aha, D. W., & Goldstone, R. L. (1992). Concept learning and flexible weighting. In *Proceedings of the 14th annual conference of the cognitive science society* (pp. 534–539). Hillsdale, NJ: Lawrence Erlbaum Associates.

Alba, J. W., & Hasher, L. (1983). Is memory schematic? *Psychological Bulletin*, *93*(2), 203–231.

Anderson, R. C., & Pichert, J. W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, *17*, 1–12.

Barsalou, L. W., & Medin, D. L. (1986). Concepts: Static dimensions or context-dependent representations? *Cahiers de Psychologie Cognitive*, *6*, 187–202.

Bower, G. H., Black, J. B., & Turner, T. J. (1979). Scripts in memory for text. *Cognitive Psychology*, *11*, 177–220.

Brewer, W. F., & Nakamura, G. V. (1984). The nature and functions of schemas. In R. S. Wyer & R. K. Srull (Eds.), *Handbook of social cognition* (Vol. 1, pp. 119–160). Hillsdale, NJ: Erlbaum.

Brewer, W. F., & Treyens, J. C. (1981). Role of schemata in memory for places. *Cognitive Psychology*, *13*(2), 207–230.

Bruce, D., & Gaines, M. T. (1976). Tests of an organizational hypothesis of isolation effects in free recall. *Journal of Verbal Learning and Verbal Behavior*, *15*, 59–72.

Busey, T. A., & Tunnicliff, J. L. (1999). Accounts of blending, distinctiveness, and typicality in the false recognition of faces. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(5), 1210–1235.

Davidenko, N., & Ramscar, M. (2004). Reexamining the "distinctiveness effect": Poorer recognition of distinctive face silhouettes. In K. D. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Conference of the Cognitive Science Society* (p. 1550). Mahwah, NJ: Lawrence Erlbaum Associates.

Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall. *Journal of Experimental Psychology*, *58*, 17–22.

Dunlosky, J., Hunt, R. R., & Clark, E. (2000). Is perceptual salience needed in explanations of the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 649–657.

Estes, W. K. (1994). *Classification and cognition.* New York: Oxford University Press.

Fabiani, M., & Donchin, E. (1995). Encoding processes and memory organization: A model of the von Restorff effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 224–240.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science*, *146*, 668–670.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology: General*, *108*, 316–355.

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 3–28.

Goodman, G. S. (1980). Picture memory: How the action schema affects retention. *Cognitive Psychology*, *12*, 473–495.

Graesser, A. C. (1981). *Prose comprehension beyond the word.* New York: Springer-Verlag.

Graesser, A. C., & Nakamura, G. V. (1982). The impact of a schema on comprehension and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 16, pp. 59–109). New York: Academic Press.

Green, R. T. (1956). Surprise as a factor in the von Restorff effect. *Journal of Experimental Psychology*, *52*, 340–344.

Hastie, R. (1981). Schematic principles in human memory. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium* (Vol. 1, pp. 39–88). Hillsdale, NJ: Erlbaum.

Hastie, R., & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology*, *37*, 25–38.

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology*, *21*, 107–112.

Hunt, R. R., & Lamb, C. A. (2001). What causes the isolation effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(6), 1359–1366.

Jenkins, W. O., & Postman, L. (1948). Isolation and 'spread of effect' in serial learning. *American Journal of Psychology*, *61*, 214–221.

Kishiyama, M. M., Yonelinas, A. P., & Lazzara, M. M. (2004). The von Restorff in amnesia: The contribution of the hippocampal system to novelty-related memory enhancements. *Journal of Cognitive Neuroscience*, *16*, 15–23.

Koffka, K. (1935). *Principles of Gestalt Psychology.* New York: Harcourt, Brace.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*, 22–44.

Langlois, J. H., Roggman, L. A., & Musselman, L. (1994). What is average and what is not average about attractive faces? *Psychological Science*, *5*(4), 214–220.

Lewandowsky, S., Kalish, M., & Ngang, S. K. (2002). Simplified learning in complex situations: Knowledge partitioning in function learning. *Journal of Experimental Psychology: General*, *131*, 163–193.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance*, *4*, 565–572.

Love, B. C. (2002). Two systems or just one? *J. S. McDonnell Foundation Cognitive Neuroscience of Category Learning Consortium Workshop.*

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of human category learning. *Psychological Review*, *111*, 309–332.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, *53*, 49–70.

Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin*, *129*, 592–613.

McArthur, L. Z. (1981). What grabs you? The role of attention in impression formation and causal attribution. In E. T. Higgins, C. P. Herman, & M. P.

Zanna (Eds.), *Social cognition: The Ontario symposium* (Vol. 1, pp. 201–246). Hillsdale, NJ: Erlbaum.

McCall, R. B., & Carriger, M. S. (1993). A meta-analysis of infant habituation and recognition memory performance as predictors of later IQ. *Child Development*, *64*(1), 57–59.

Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review*, *85*, 207–238.

Metcalfe, J. (1993). Novelty monitoring, metacognition, and control in a composite holographic associative recall model: Implications for Korsakoff amnesia. *Psychological Review*, *100*(1), 3–22.

Murphy, G. L., & Allopenna, P. D. (1994). The locus of knowledge effects in concept learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 904–919.

Nosofsky, R. M. (1986). Attention, similarity, and the identification-categorization relationship. *Journal of Experimental Psychology: General*, *115*, 39–57.

Nosofsky, R. M. (1988). Similarity, frequency, and category representations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 54–65.

Nosofsky, R. M., Palmeri, T. J., & McKinley, S. C. (1994). Rule-plus-exception model of classification learning. *Psychological Review*, *101*(1), 53–79.

Nosofsky, R. M., & Zaki, S. F. (2003). A hybrid-similarity exemplar model for predicting distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1194–1209.

Ono, H. (1967). Difference threshold for stimulus length under simultaneous and nonsimultaneous viewing conditions. *Perception & Psychophysics*, *2*, 201–207.

Palmeri, T. J., & Nosofsky, R. M. (1995). Recognition memory for exceptions to the category rule. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 548–568.

Pazzani, M. J. (1991). Influence of prior knowledge on concept acquisition: Experimental and computational results. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*, 416–432.

Pezdek, K., Whetstone, T., Reynolds, K., Askari, N., & Dougherty, T. (1989). Memory for real-world scenes: The role of consistency with schema expectation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*(4), 587–595.

Pichert, J. W., & Anderson, R. C. (1977). Taking different perspectives on a story. *Journal of Educational Psychology*, *69*, 309–315.

Pinker, S. (1991). Rules of language. *Science*, *253*, 530–535.

Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology*, *77*, 353–363.

Ranganath, C., & Rainer, G. (2003). Neural mechanisms for detecting and remembering novel events. *Nature Reviews: Neuroscience*, *4*, 193–202.

Roediger, H. L. I., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814.

Rojahn, K., & Pettigrew, T. F. (1992). Memory for schema-relevant information: A meta-analytic resolution. *British Journal of Social Psychology*, *31*(2), 81–109.

Rolf, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality & Social Psychology Review*, *8*(4), 364–382.

Rothbart, M., Evans, M., & Fulero, S. (1979). Recall for confirming events: Memory processes and the maintenance of social stereotypes. *Journal of Experimental Social Psychology*, *15*, 343–355.

Sakamoto, Y., & Love, B. C. (2004). Schematic influences on category learning and recognition memory. *Journal of Experimental Psychology: General*, *133*, 534–553.

Sakamoto, Y., Matsuka, T., & Love, B. C. (2004). Dimension-wide vs. exemplar-specific attention in category learning and recognition. In M. Lovett, C. Schunn, C. Lebiere, & P. Munro (Eds.), *Proceedings of the 6th International Conference of Cognitive Modeling* (pp. 261–266). Mahwah, NJ: Lawrence Erlbaum Associates.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding.* Hillsdale, NJ: Erlbaum.

Schmidt, S. R. (1991). Can we have a distinctive theory of memory? *Memory & Cognition*, *19*(6), 523–542.

Schyns, P. G., & Murphy, G. L. (1994). The ontogeny of part representation in object concepts. In D. L. Medin (Ed.), *The psychology of learning and motivation* (Vol. 31, pp. 305–354). San Diego, CA: Academic Press.

Sentis, K. P., & Burnstein, E. (1979). Remembering schema-consistent information: Effects of a balance schema on recognition memory. *Journal of Personality and Social Psychology*, *37*(12), 2200–2211.

Sherman, J. W., & Frost, L. A. (2000). On the encoding of stereotype-relevant information under cognitive load. *Personality and Social Psychology Bulletin*, *26*(1), 26–34.

Shiffrin, R. M., Huber, D. E., & Marinelli, K. (1995). Effects of category length and strength on familiarity in recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(2), 267–287.

Snyder, M., & Uranowitz, S. W. (1978). Reconstructing the past: Some cognitive consequences of person perception. *Journal of Personality and Social Psychology*, *36*, 941–950.

Srull, T. K. (1981). Person memory: Some tests of associative storage and retrieval models. *Journal of Experimental Psychology: Human Learning and Memory*, *7*, 440–463.

Srull, T. K., & Wyer, R. S. J. (1989). Person memory and judgment. *Psychological Review*, *96*, 58–83.

Stangor, C., & McMillan, D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: A review of the social and developmental literatures. *Psychological Bulletin*, *111*, 42–61.

Taylor, S. E., & Crocker, J. (1981). Schematic bases of social information processing. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario symposium* (Vol. 1, pp. 89–134). Hillsdale, NJ: Erlbaum.

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*, 778–793.

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *The Quarterly Journal of Experimental Psychology*, *43A*, 161–204.

von Restorff, H. (1933). Analyse von vorgängen in spurenfeld. I. Über die wirkung von bereichsbildungen im spurenfeld [Analysis of processes in the memory trace. I. On the effect of group formations on the memory trace]. *Psychologische Forschung*, *18*, 299–342.

Wallace, W. P. (1965). Review of the historical, empirical, and theoretical status of the von Restorff phenomenon. *Psychological Bulletin*, *63*, 410–424.

Wattenmaker, W. D., Dewey, G. I., Murphy, T. D., & Medin, D. L. (1986). Linear separability and concept learning: Context, relational properties, and concept naturalness. *Cognitive Psychology*, *18*, 158–194.

Wisniewski, E. J., & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, *18*, 221–281.

Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language*, *39*, 124–149.

Zaki, S. R., & Nosofsky, R. M. (2001). Exemplar accounts of blending and distinctiveness effects in perceptual old-new recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(4), 1022–1041.

Vita

Yasuaki Sakamoto was born in Japan on September 15, 1974, the son of Shinichi Sakamoto and Yoko Sakamoto. He received the degree of Bachelor of Arts from Indiana University at Bloomington in 1999. During the following year he was employed as a research assistant in the Percepts and Concepts laboratory at Indiana University. In September, 2000, he entered The Graduate School at The University of Texas at Austin. In December, 2003, he received the degree of Master of Arts from The University of Texas at Austin.

Permanent Address: 3024 B Windsor Rd.
Austin, TX 78703

This dissertation was typeset with LaTeX 2_ε[3] by the author.

[3] LaTeX 2_ε is an extension of LaTeX. LaTeX is a collection of macros for TeX. TeX is a trademark of the American Mathematical Society.